Displacement Convexity – A Useful Framework for the Study of Spatially Coupled Codes

Rafah El-Khatib
EPFL, Switzerland
Email: rafah.el-khatib@epfl.ch

Nicolas Macris
EPFL, Switzerland
Email: nicolas.macris@epfl.ch

Ruediger Urbanke
EPFL, Switzerland
Email: ruediger.urbanke@epfl.ch

***Abstract*—Spatial coupling has recently emerged as a powerful paradigm to construct graphical models that work well under low-complexity message-passing algorithms. Although much progress has been made on the analysis of spatially coupled models under message passing, there is still room for improvement, both in terms of simplifying some existing proofs as well as in terms of proving some additional properties.**

We introduce one further tool for the analysis, namely the concept of displacement convexity. This concept plays a crucial role in the theory of optimal transport and it is also very well suited for the analysis of spatially coupled systems. In cases where the concept applies, displacement convexity allows functionals of distributions which are not convex to be represented in an alternative form, so that they are convex with respect to the new parameterization. The additional convex structure can then often be used to prove the uniqueness of the minimizer of this functional. As a proof of concept we consider spatially coupled (l, r)-regular Gallager ensembles when transmission takes place over the binary erasure channel. In particular, we prove that the potential functional governing this system admits a unique optimal "profile" which characterizes the "decoding wave" of the spatially coupled system using the tool of displacement convexity.

I. Introduction

Spatially coupled codes were introduced in the form of low-density parity-check codes by Felstrom and Zigangirov in [1]. Such codes are constructed by spatially coupling nearby replicas of a code defined on a graph. It has been proven that such ensembles perform very well under low-complexity message-passing algorithms. Indeed, this combination achieves essentially optimal performance. More generally, the concept of spatial coupling is proving to be very useful not only for coding but also in statistical physics and compressive sensing. Given this range of applications, it is worth investigating basic properties of this construction in generality. Our aim is to introduce one further tool for the analysis of such systems – namely the concept of displacement convexity. Displacement convexity plays a crucial role in the theory of optimal transport. But it is also very well suited as a tool for the analysis of spatially coupled graphical models.

Using the simple case of regular LDPC ensembles and transmission over the binary erasure channel (BEC), we explain how the concept of displacement convexity can help to simplify some existing proofs and to derive new results. The actual range of applications which can benefit from this concept is considerably larger and we pose some open problems in this respect as further research directions.

One of the most important properties of spatially coupled codes is that they exhibit the so-called threshold saturation phenomenon. That is, spatially coupled ensembles generically have a BP threshold which is as large as the maximum-a posteriori (MAP) threshold of the underlying ensemble, i.e., their threshold has *saturated* to the largest possible value. This result has been proved for transmission over the BEC in [2], [3], and [4] and for transmission over general binary-input memoryless output-symmetric (BMS) channels in [5] and [6].

Our goal is to provide a new proof technique that targets the same problem. The tool we introduce is applicable to systems which are governed by a variational principle; we use (l, r)-regular Gallager ensembles as a proof of concept. We run the belief-propagation algorithm on this ensemble and express the potential functional using the results in [3]. The potential functional is the analog of the Bethe free energy, and the density evolution (DE) equations can be obtained by differentiating this potential. Our goal is to find the fixed-point (FP) solution(s) of the DE equations, which is equivalent to finding the minimizer(s) of the potential.

For the simple case we analyze, we use displacement convexity [7] to prove that the potential describing the system is convex with respect to an alternative structure of probability measures. We consider the static case, when the decoder phase transition threshold is equal to the maximum a posteriori (MAP) threshold. Using this alternative convex structure, we then prove that the potential of the ensemble admits a unique minimizing profile.

The paper is organized as follows: Section II introduces the framework for our analysis and our main results. We then give a quick introduction of the notion of displacement convexity in Section III. Finally, Section IV presents a proof of existence of the profile that minimizes the potential, and Section V proves that this minimizer is unique.

II. Setting and Main Results

In this section, we introduce the model and the associated variational problem to which we apply the displacement convexity proof technique, and we state our main result.

A. (l, r, L, w)-Regular Ensembles on the BEC

Consider the spatially coupled (l, r, L, w)-regular ensemble, described in detail in [5], where the parameters represent the left degree, right degree, system length parameter, and coupling window size (or smoothing parameter), respectively. Specifically, the ensemble is constructed as follows: consider

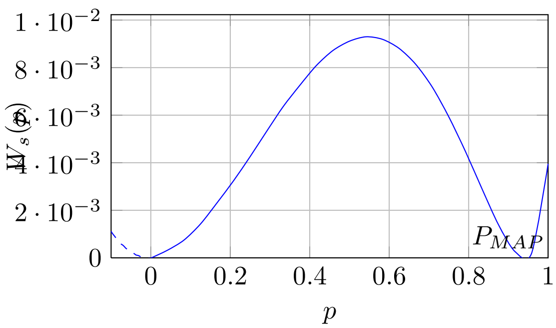

Fig. 1. The plot of the single system potential $W_s(p)$ as a function of the check-erasure probability p, for a $(3,6)$ uncoupled ensemble and $\epsilon = \epsilon_{\rm MAP}$. There are two minima at $p=0$ and $p = p_{\rm MAP}$.

$2L+1$ replicas of a protograph of an (l,r)-regular ensemble. We couple these components by connecting every variable node to l check nodes, and every check node to r variable nodes. The connections are chosen randomly: for a variable node at position z, each of its l connections is chosen uniformly and independently in the range $[z,\dots,z+w-1]$, and for a check node at position z, each of its r connections is chosen uniformly and independently in the range $[z-w+1,\dots,z]$.

For the channel we take a BEC with parameter ϵ. If $x_z;\ z\in[-L,\dots,L]$ denotes the erasure probability of the variable node at position z, the fixed-point (FP) condition implied by density evolution is

$$x_z = \epsilon\Big(1-\frac{1}{w}\sum_{i=0}^{w-1}(1-\frac{1}{w}\sum_{j=0}^{w-1}x_{z+i-j})^{r-1}\Big)^{l-1}.$$

This FP condition can be obtained by minimizing a "potential functional". It can be shown using the results in [3] that the potential is

$$\frac{1}{w}\sum_{z=-L}^{L}\Big\{-x_z(1-x_z)^{r-1}+\frac{1}{r}-\frac{1}{r}(1-x_z)^r \\ -\frac{\epsilon}{l}\Big[\frac{1}{w}\sum_{u=0}^{w-1}[1-(1-x_{z+u})^{r-1}]\Big]^l\Big\}. \tag{1}$$

At this point the normalization $1/w$ is a convenience whose reason will immediately appear.

The natural setting for displacement convexity is the continuum case. We will therefore consider the continuum limit of (1). Extending our results to the discrete setting is one among various open problems to which we briefly come back in the conclusion. We define the rescaled variables $\tilde{z}=\frac{z}{w}$, $\tilde{u}=\frac{u}{w}$ and the rescaled function $\tilde{x}(\frac{z}{w})\equiv x_z$. It is easy to see that (1) becomes a Riemann sum. When we take the limit $L\to+\infty$ first and then $w\to+\infty$, we find

$$\int_{\mathbb{R}} \mathrm{d}\tilde{z}\Big\{-\tilde{x}(\tilde{z})(1-\tilde{x}(\tilde{z}))^{r-1}-\frac{1}{r}(1-\tilde{x}(\tilde{z}))^r \\ +\frac{1}{r}-\frac{\epsilon}{l}\Big[\int_0^1 \mathrm{d}\tilde{u}\,[1-(1-\tilde{x}(\tilde{z}+\tilde{u}))^{r-1}]\Big]^l\Big\}. \tag{2}$$

At this point the reader might wonder if the integrals converge. As explained in the introduction, we look in this paper at the decoder phase transition threshold $\epsilon=\epsilon_{\rm MAP}$. We give at the end of this paragraph the conditions on the erasure probability profile needed to have a well defined problem. From now on, the reader should think of the noise level as fixed to the value $\epsilon=\epsilon_{\rm MAP}$, although we abuse notation by simply writing ϵ in the formulas that follow.

It is more convenient to express (2) with the function $p(z)=1-(1-\tilde{x}(z))^{r-1}$. Note that this function is interpreted as the erasure probability emitted by check nodes. Summarizing, the potential functional of interest is

$$\mathcal{W}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z\Big\{\Big(1-\frac{1}{r}\Big)(1-p(z))^{\frac{r}{r-1}}-(1-p(z)) \\ +\frac{1}{r}-\frac{\epsilon}{l}\Big(\int_0^1 \mathrm{d}u\,\Big[1-(1-m(z+u))^{r-1}\Big]\Big)^l\Big\}. \tag{3}$$

A word about the notation here: we use *square* brackets for functionals, i.e. "functions of functions" and usual *round* brackets for functions of a real variable. Equation (3) can be expressed as a sum of two contributions $\mathcal{W}_{\rm single}[p(\cdot)]+\mathcal{W}_{\rm int}[p(\cdot)]$ which are defined as follows:

$$\mathcal{W}_{\rm single}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z\Big\{\Big(1-\frac{1}{r}\Big)(1-p(z))^{\frac{r}{r-1}}-(1-p(z)) \\ +\frac{1}{r}-\frac{\epsilon}{l}p(z)^l\Big\} \equiv \int_{\mathbb{R}} \mathrm{d}z\,W_s(p(z)), \tag{4}$$

$$\mathcal{W}_{\rm int}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z\,\frac{\epsilon}{l}\Big\{p(z)^l-\Big[\int_0^1 \mathrm{d}u\,p(z+u)\Big]^l\Big\}. \tag{5}$$

We call (4) the "single system potential functional" and (5) the "interaction functional". The following remarks explain the interpretation suggested by these names. The term (5) vanishes when evaluated for a constant $p(z)=p$. Moreover the integrand of (4), namely $W_s(p(z))=W_s(p)$ is just the potential of the underlying uncoupled code ensemble. This is easily seen by recognizing that the usual density evolution equation for the erasure probability of checks is recovered by setting the derivative of $W_s(p)$ to zero. We will call $W_s(p)$ the "single system potential". A plot of $W_s(p)$ for the $(3,6)$-ensemble is shown as an example in Figure I when $\epsilon=\epsilon_{\rm MAP}$. The figure shows that the single potential vanishes at $p=0$ and $p=p_{\rm MAP}$, some positive value. This is a generic feature of all (l,r)-regular code ensembles as long as $l\geq 3$ (for cycle codes $l=2$, we have $p_{\rm MAP}=0$). This shows, in particular, that in order for the integrals in (3) to be well defined, we have to consider profiles $p(\cdot)\in\mathcal{S}$ where

$$\mathcal{S}=\{p(\cdot):\mathbb{R}\to[0,p_{\rm MAP}]\ \text{s.t}\lim_{z\to-\infty}zp(z)=0, \\ \lim_{z\to+\infty}z(p(z)-p_{\rm MAP})=0\}. \tag{6}$$

In particular the left limit is 0 and the right limit is $p_{\rm MAP}$.

B. Main Results

It is easy to see that $\mathcal{W}[p(\cdot)]$ is bounded from below, more precisely $\inf_{p(\cdot)\in\mathcal{S}} \mathcal{W}[p(\cdot)] \geq -\frac{1}{2}p_{\text{MAP}}^l$. Indeed $W_s(p) \geq 0$ as seen in Figure I and using Jensen one can show $\mathcal{W}_{\text{int}}[p(\cdot)] \geq -\frac{1}{2}p_{\text{MAP}}^l$ (see Lemma 4.1). The first non-trivial question one may ask is whether a minimum is attained in $\mathcal{S}$. This question can be answered by means of the so-called "direct method", a standard method of calculus of variations [13], [14], (see Section IV-B). The next natural question is whether there is a unique minimizer. The functional is translation invariant i.e., $\mathcal{W}[p(\cdot + a)] = \mathcal{W}[p(\cdot)]$ so it is natural to ask whether there is a unique minimizer up to translation. When a functional is strictly convex the answer is immediate. Unfortunately, $\mathcal{W}[p(\cdot)]$ is *not convex* (and hence not strictly convex). However, and this is the main point of this paper, we will show that it can be represented in an alternative form so that convexity (and strict convexity) can be used in this alternative domain. This alternative notion of convexity that we will use is briefly explained in the next section. It will allow us to prove:

Theorem 2.1: Let $\epsilon = \epsilon_{\text{MAP}}$. The functional $\mathcal{W}[p(\cdot)]$ attains its minimum on $\mathcal{S}$. Moreover, this minimum is unique up to translations.

Let us briefly outline the proof strategy of this theorem. We show using rearrangement inequalities [12] that the minimizers in $\mathcal{S}$ are necessarily increasing (see Lemma 4.3). Note that increasing profiles can be viewed as cumulative distribution functions (cdfs) and in particular they have an inverse function. This turns out to be the right setting both for the existence proof (Section IV-B) and for the application of the displacement convexity technique to prove uniqueness (Section V).

III. Displacement Convexity

Displacement convexity can be very useful in functional analysis. It goes back to McCann [7] and plays an important role in the theory of optimal transport [8]. It has been used in [9] and [10] to study a functional governing a spatially coupled Curie-Weiss model [11], which bears close similarities with the coding theory model studied here. In this section, we give a quick introduction to the tool of displacement convexity.

Recall first that the usual notion of convexity of a generic functional $\mathcal{F}[p(\cdot)]$ on a generic space $\mathcal{X}$ means that for all $p(\cdot), p'(\cdot) \in \mathcal{X}$

$$\mathcal{F}[p_\lambda(\cdot)] \leq \lambda\mathcal{F}[p(\cdot)] + (1-\lambda)\mathcal{F}[p'(\cdot)],$$

where $p_\lambda(z) = (1-\lambda)p(z) + \lambda p'(z)$, $\lambda \in [0,1]$.

Lemma 4.3 in Section IV shows that we can restrict the minimization problem to the space of increasing profiles. Thus, the discussion below assumes that we consider only such profiles. This is the correct setting for defining displacement convexity.

An increasing profile with left limit 0 and right limit p_{MAP} can be thought of as a cdf (up to scaling because the right limit is not 1). Further, such increasing functions have increasing inverse functions (which can also be thought of as cdfs, up to scaling). More precisely, consider the following bijective maps that associate (with an abuse of notation) to a cdf $p(\cdot)$ its inverse $z(\cdot)$:

$$z(p) = \inf\{\, z : \ p(z) > p\},$$
$$p(z) = \inf\{\, p : \ z(p) > z\}.$$

For any two increasing profiles $p(\cdot), p'(\cdot) \in \mathcal{S}$, we consider $z(\cdot), z'(\cdot)$ their respective inverses under the maps defined above. Then for any $\lambda \in [0,1]$, the interpolated profile $p_\lambda(\cdot)$ is defined as follows:

$$z_\lambda(p) = (1-\lambda)z(p) + \lambda z'(p),$$
$$p_\lambda(z) = \inf\{\, p : \ z_\lambda > z\}.$$

In words, the difference in interpolation under the alternative structure is that the linear interpolation is applied on the inverse of the profiles of interest, and the effect of such an interpolation is then mapped back into the space of profiles. Displacement convexity of $\mathcal{W}[p(\cdot)]$ on the space $\mathcal{S}$ simply means that the following inequality holds:

$$\mathcal{W}[p_\lambda(\cdot)] \leq (1-\lambda)\mathcal{W}[p(\cdot)] + \lambda\mathcal{W}[p'(\cdot)] \tag{7}$$

for any $p(\cdot), p'(\cdot) \in \mathcal{S}$. Strict displacement convexity means that this inequality is strict as long as $p(\cdot)$ and $p'(\cdot)$ are distinct. We will prove the strict displacement convexity of $\mathcal{W}[p(\cdot)]$ by separately proving this property, in Sections V-A and V-B, for the two functionals (4) and (5), respectively.

IV. Existence of Minimizing Profile

In this section, we prove that the functional $\mathcal{W}$ attains its minimum.

A. Preliminaries

We start by some preliminaries to show that one can restrict the search of minimizing profiles to those in $\mathcal{S}$ that are monotone increasing. The proofs of the lemmas can be found in Appendices A-D.

The first lemma states that the interaction potential is bounded from below.

Lemma 4.1: For any $p(\cdot)$ in $\mathcal{S}$,

$$\int_{\mathbb{R}} \mathrm{d}z \left\{ p(z)^l - \left(\int_0^1 \mathrm{d}u \, p(z+u) \right)^l \right\} \geq -\frac{1}{2}p_{\text{MAP}}^l.$$

Proof: See Appendix A. ■

We remark that any constant lower bound is sufficient for our purposes: finding profiles that minimize a potential is equivalent to find those that minimize a potential added to a constant. The following lemma states that a truncation of the profile at the value p_{MAP} decreases the potential functional, so we may restrict our search of minimizing profiles to those with range $p(z) \in [0, p_{\text{MAP}}]$.

Lemma 4.2: Define $\bar{p}(z) = \min\{p(z), p_{\text{MAP}}\}$. For all $p \in \mathcal{S}$ we have

$$\mathcal{W}[p(\cdot)] \geq \mathcal{W}[\bar{p}(\cdot)].$$

Proof: See Appendix B. ■

We next restrict our search of minimizing profiles to increasing ones. In order to achieve this we will use rearrangement

inequalities. Let us first recall the notion of an increasing rearrangement. In words, an increasing rearrangement associates to any function $p(\cdot) \in \mathcal{S}$ with range $[0, p_{\text{MAP}}]$ an increasing function $p^*(\cdot)$ so that the total mass is preserved. More formally, any non-negative function can be represented in layer cake form

$$p(z) = \int_0^{+\infty} dt\, \mathbb{1}_{E_t}(z),$$

where $\mathbb{1}_{E_t}(z)$ is the indicator function of the level set $E_t = \{z | p(z) > t\}$. For each t, the level set E_t can be written as the union of a bounded set A_t and a half line $]a_t, +\infty[$. We define the rearranged set $E_t^* =]a_t - |A_t|, +\infty[$. The increasing rearrangement of $p(\cdot)$ is the new function $p^*(\cdot)$ whose level sets are E_t^*. More explicitly,

$$p^*(z) = \int_0^{+\infty} dt\, \mathbb{1}_{E_t^*}(z).$$

Lemma 4.3: Take any $p(\cdot) \in \mathcal{S}$ and let $p^*(\cdot)$ be its increasing rearrangement. Then,

$$\mathcal{W}[p(\cdot)] \geq \mathcal{W}[p^*(\cdot)].$$

Proof: See Appendix C. ■

We can thus restrict the search of minimizing profiles to the space of increasing profiles. Furthermore because of translation invariance of the potential functional it is always possible to translate a increasing profile so that $p(z)$ crosses the value $\frac{p_{\text{MAP}}}{2}$ at the origin. In other words $\inf_{\mathcal{S}} \mathcal{W}[p(\cdot)] \geq \inf_{\mathcal{S}'} \mathcal{W}[p(\cdot)]$ with

$$\mathcal{S}' = \{p(\cdot) \in \mathcal{S} : \text{ increasing and } , p(z) \leq \frac{p_{\text{MAP}}}{2}, z \leq 0, \\ p(z) > \frac{p_{\text{MAP}}}{2}, z > 0\}.$$

In particular, we can think of a profile as a cdf associated to an underlying measure (up to scaling).

The final step of these preliminaries concerns a necessary condition that any *minimizing sequence* must satisfy. A minimizing sequence in $\mathcal{S}'$ is by definition any sequence $p_n(\cdot) \in \mathcal{S}'$ such that

$$\lim_{n\to\infty} \mathcal{W}[p_n(\cdot)] = \inf_{p \in \mathcal{S}'} \mathcal{W}[p(\cdot)]. \tag{8}$$

Such a sequence exists as long as the functional is bounded from below; which is indeed true because of Lemma 4.1 and $W_s(p) \geq 0$. Consider the sequence of probability measures associated to the sequence of cdfs $p_n(\cdot)$. The following lemma states that this sequence of measures is *tight*.

Lemma 4.4: Let $p_n(\cdot) \in \mathcal{S}'$ be a minimizing sequence of cdfs. For any $\delta > 0$ we can find $M_\delta > 0$ (independent of n) such that

$$p_n(M_\delta) - p_n(-M_\delta) > (1 - \delta) p_{\text{MAP}}$$

for all n.

Proof: See Appendix D. ■

B. The Direct Method

The direct method in the calculus of variations [13]-[14] is a standard scheme to prove that minimizers exist. We use this method to obtain the following theorem:

Theorem 4.5: The potential functional $\mathcal{W}[p(\cdot)]$ attains its minimum in the space $\mathcal{S}'$, and hence in $\mathcal{S}$.

Proof: Let us take any minimizing sequence $p_n(\cdot)$ of cdfs, i.e. a sequence that satisfies (8). By Lemma 4.4 the corresponding sequence of measures is tight. Thus by a simple version of Prokhorov's theorem for measures on the real line, we can extract a (point-wise) convergent subsequence of cdfs $p_{n_k}(\cdot) \to p_0(\cdot)$ as $k \to +\infty$ with $p_0(\cdot) \in \mathcal{S}'$. By Fatou's Lemma, one can check that the potential functional is lower-semi-continuous, which means

$$\mathcal{W}[p_0(\cdot)] \leq \liminf_{k\to+\infty} \mathcal{W}[p_{n_k}(\cdot)]. \tag{9}$$

Putting (8) and (9) together,

$$\mathcal{W}[p_0(\cdot)] \leq \liminf_{k\to+\infty} \mathcal{W}[p_{n_k}(\cdot)] \leq \lim_{n\to+\infty} \mathcal{W}[p_n(\cdot)] = \inf_{\mathcal{S}'} \mathcal{W}[p(\cdot)].$$

On the other hand $\inf_{\mathcal{S}'} \mathcal{W}[p(\cdot)] \leq \mathcal{W}[p_0(\cdot)]$. Thus we conclude $\inf_{\mathcal{S}'} \mathcal{W}[p(\cdot)] = \mathcal{W}[p_0(\cdot)]$. ■

V. Proof of Displacement Convexity of the Functional

This section contains the main result of the paper, namely that the potential functional $\mathcal{W}[p(\cdot)]$ is strictly displacement convex in $\mathcal{S}'$. Concretely, we prove the strict form of (7). This allows to conclude that the minimizer is unique.[1] Hence we also conclude that the DE equations governing the system have a unique FP solution.

A. Displacement Convexity of the Single-Potential Term

We first prove that the single-potential functional $\mathcal{W}_{\text{single}}[p(\cdot)]$ is displacement convex. The proof is very simple. Note the single system potential $W_s(p)$ is *not convex* in the usual sense (see Figure I).

Theorem 5.1: Let $p(\cdot)$ and $p'(\cdot)$ be in $\mathcal{S}'$ and let $p_\lambda(\cdot)$ the interpolating profile as defined in Section III. Then

$$\mathcal{W}_{\text{single}}[p_\lambda(\cdot)] = (1 - \lambda)\mathcal{W}_{\text{single}}[p(\cdot)] + \lambda \mathcal{W}_{\text{single}}[p'(\cdot)].$$

Proof: Recall that $\mathcal{W}_{\text{single}}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z W_s(p(z))$. Recall also that $p_\lambda(z)$ as defined in Section III is the inverse of $z_\lambda(p) = \lambda z(p) + (1 - \lambda) z'(p)$. Thus

$$\begin{aligned} \int_{\mathbb{R}} \mathrm{d}z W_s(p_\lambda(z)) &= \int_0^{p_{\text{MAP}}} \mathrm{d}z_\lambda(p) W_s(p) \\ &= (1-\lambda) \int_0^{p_{\text{MAP}}} \mathrm{d}z(p) W_s(p) + \lambda \int_0^{p_{\text{MAP}}} \mathrm{d}z'(p) W_s(p) \\ &= (1-\lambda) \int_{\mathbb{R}} \mathrm{d}z W_s(p(z)) + \lambda \int_{\mathbb{R}} \mathrm{d}z W_s(p'(z)). \end{aligned}$$

■

Thus, the function $\lambda \to \mathcal{W}_{\text{single}}[p_\lambda(\cdot)]$ is linear, hence convex.

[1] Recall that in the definition of $\mathcal{S}'$ we have removed the translational degree of freedom.

B. Displacement Convexity of the Interaction-Potential Term

The study of strict displacement convexity of the interaction potential term involves a more complicated analysis.

Theorem 5.2: Let $p(\cdot)$ and $p'(\cdot)$ be in $\mathcal{S}'$ and let $p_\lambda(\cdot)$ the interpolating profile as defined in Section III. Then

$$\mathcal{W}_{\rm int}[p_\lambda(\cdot)] < (1-\lambda)\mathcal{W}_{\rm int}[p(\cdot)] + \lambda\mathcal{W}_{\rm int}[p'(\cdot)]. \qquad (10)$$

Proof: Since p can be seen as a cdf we associate with it its probability measure μ such that $p(z) = p_{\rm MAP}\int_{-\infty}^{z} \mathrm{d}\mu(x)$. Let us rewrite the interaction functional in the form

$$\mathcal{W}_{\rm int}[p_\lambda(\cdot)] = \int_{\mathbb{R}} \mathrm{d}\mu_\lambda(x_1)\dots\mathrm{d}\mu_\lambda(x_l)V(x_1,\dots,x_l), \quad (11)$$

where $V(x_1,\dots,x_l)$ is a totally symmetric "kernel function" that we will compute. There is an argument (see [8]) that allows to conclude (10) whenever V is jointly convex (in the usual sense). Let us briefly explain this argument here. Consider the measures μ, μ' associated to cdfs $p(\cdot)$, $p'(\cdot)$. Then there exists a unique increasing map $T:\mathbb{R}\to\mathbb{R}$ such that $\mu' = T\#\mu$. Here $T\#\mu$ is the push-forward[2] of μ under T. Then from $x_\lambda(p) = (1-\lambda)x(p) + \lambda x'(p)$ we have that $\mu_\lambda = T_\lambda\#\mu$ where $T_\lambda(x) = \lambda x + (1-\lambda)T(x)$. Equation (11) can be written as

$$\begin{aligned}\mathcal{W}_{\rm int}[p_\lambda(\cdot)] &= \int_{\mathbb{R}} \mathrm{d}\mu(x_1)\dots\mathrm{d}\mu(x_l)V(T_\lambda(x_1),\dots,T_\lambda(x_l))\\ &= l!\int_{S_x} \mathrm{d}\mu(x_1)\dots\mathrm{d}\mu(x_l)V(T_\lambda(x_1),\dots,T_\lambda(x_l)).\end{aligned}$$

In the second equality we restrict the integrals over the sector $S_x = \{\mathbf{x} = (x_1,\cdots,x_l) : x_i \geq x_j \text{ if } i<j\}$, which is possible since V is totally symmetric. Now it is important to notice that since T is an increasing map we have $T_\lambda(x_1)\geq\cdots\geq T_\lambda(x_l)$ for any $\lambda\in[0,1]$. Moreover the λ dependence in the kernel function is linear. Thus the proof of displacement convexity ultimately rests on checking that the kernel function is jointly convex in one sector, say S_x.[3] In fact the kernel function is translation invariant and can be expressed as a function of the distances $d_{1i}\equiv x_1 - x_i$, $i=1,\dots,l$. We will prove joint convexity of V as a function of these distances. To get strict displacement convexity on $\mathcal{S}'$ we require convexity of the kernel function, and *strict* convexity as a function of $\mathbf{d} = (d_{12},\dots,d_{1l})$ on a *non-zero measure subset* of the $\mathbf{d}$-space.

Now it remains to compute V and to investigate its convexity. With appropriate usage of Fubini's theorem and after some computations, we find

$$\mathcal{W}_{\rm int}[p(\cdot)] = \frac{\epsilon p_{\rm MAP}^l}{l}\int_{\mathbb{R}^l}\prod_{i=1}^{l}\mathrm{d}\mu(x_i)\Bigg\{\int_{[0,1]^l}\prod_{i=1}^{l}\mathrm{d}u_i \int_{\mathbb{R}}\mathrm{d}z\Bigg(\prod_{i=1}^{l}\theta(z-x_i) - \prod_{i=1}^{l}\theta(z-(x_i-u_i))\Bigg)\Bigg\}, \quad (12)$$

where $\theta(x)$ denotes the heaviside step function at x. So the kernel $V(x_1,\dots,x_l)$ in (11) is the integrand of the first l integrals in (12). Our goal henceforth is to prove that V is convex in the usual sense. We will prove a stronger statement, namely that $V_{\mathbf{u}}$ is convex, where $\mathbf{u} = (u_1,\dots,u_l)$,

$$V_{\mathbf{u}}(\mathbf{x}) = \int_{\mathbb{R}} dz\Bigg(\prod_{i=1}^{l}\theta(z-x_i) - \prod_{i=1}^{l}\theta(z-(x_i-u_i))\Bigg).$$

We recall here that we restrict our analysis to the sector of the space of variables S_x. Also, we remark that $\prod_{i=1}^{l}\theta(a_i) = \theta(\max_{i=1\dots l} a_i)$. We observe that $V_{\mathbf{u}}$ can be written in terms of the distances $d_{1i} = x_1 - x_i$, $i=2,\dots,l$ as (here $d_{1i}\equiv 0$)

$$\begin{aligned}V_{\mathbf{u}}(\mathbf{x}) &= \int_{\mathbb{R}}\mathrm{d}z\{\theta(z-x_1) - \theta(\max_{i=1\dots l}(z-(x_i-u_i)))\}\\ &= -\min_{i=1\dots l}(x_1 - x_i + u_i) = -\min_{i=1\dots l}(d_{1i}+u_i)\end{aligned}$$

Lemma 5.3 below states that $V_{\mathbf{u}}(\mathbf{x})$ is jointly convex in S_x for all $\mathbf{u}$. This implies that $V(\mathbf{x})$ is jointly convex in S_x. In Appendix E we give explicit formulas for V in terms of the distances d_{1i} and prove that it is strictly convex in a subset of non-zero measure. Concretely, this subset is a small neighborhood of the origin $d_{1i}=0$, $i=2,\dots,l$. This completes the proof. ∎

Lemma 5.3: The function $f_{\mathbf{u}}(\mathbf{d}) = \min_i(d_{1i}+u_i)$ is concave in $\mathbf{d}$, where $\mathbf{d} = (d_{12},\dots,d_{1l})$ and $d_{1i}\equiv 0$.

Proof: Let $\mathbf{d}$ and $\mathbf{d}'$ be two instances of the argument of $f_{\mathbf{u}}$. Then, for $\lambda\in[0,1]$,

$$\begin{aligned}f_{\mathbf{u}}((1-\lambda)\mathbf{d}+\lambda\mathbf{d}') &= \min_i((1-\lambda)d_{1i} + \lambda d'_{1i} + u_i)\\ &= \min_i((1-\lambda)(d_{1i}+u_i) + \lambda(d'_{1i}+u_i))\\ &\geq (1-\lambda)\min_i(d_{1i}+u_i) + \lambda\min_i(d'_{1i}+u_i)\\ &= (1-\lambda)f_{\mathbf{u}}(\mathbf{d}) + \lambda f_{\mathbf{u}}(\mathbf{d}').\end{aligned}$$

This shows concavity. ∎

VI. Conclusion

In this paper, we demonstrate a new tool for the analysis of spatially coupled codes, namely the concept of displacement convexity. This tool makes use of an alternative structure of probability distributions and hence applies to an appropriate space of increasing profiles. We expressed the potential functional governing the (l,r)-regular ensemble and proved that it indeed admits a minimizing profile in the appropriate space of profiles, and that it is strictly convex under the alternative structure. This result implies that the potential functional admits a unique minimizing profile, or equivalently, that the DE equations governing the system admit a unique FP solution.

There are several questions that can be posed in this context. First, we recall that the original potential functional governing the system at hand is in discrete form. Can one extend the displacement convexity framework to the discrete setting? Displacement convexity can presumably be used to analyze a large range of problems with flavors similar to the present one. In particular, it can presumably be applied to general BMS channels, the random K-SAT and Q-coloring problems to name a few. We plan to come back to these problems in the future.

[2]Given a measurable map $T:\mathbb{R}\to\mathbb{R}$, the push-forward of ν under T is the measure $T\#\nu$ such that, for any bounded continuous function ϕ, $\int_{\mathbb{R}}\phi(T(x))\mathrm{d}\nu(x) = \int_{\mathbb{R}}\phi(x)\mathrm{d}(T\#\nu)(x)$.

[3]By symmetry, convexity in one sector implies convexity in other sectors. However this does not mean that convexity holds if arguments are taken in different sectors. And, indeed in the present problem one can check that convexity only holds within each sector.

APPENDIX

A. Positivity of the Interaction Potential

Proof of Lemma 4.1: From Jensen's inequality,

$$\int_0^1 \mathrm{d}u\, p(z+u)^l \geq \Big(\int_0^1 \mathrm{d}u\, p(z+u)\Big)^l. \tag{13}$$

Further,

$$\int_{-M}^{M} \mathrm{d}z \int_0^1 \mathrm{d}u\, p(z+u)^l \overset{(a)}{=} \int_0^1 \mathrm{d}u \int_{-M+u}^{M+u} \mathrm{d}z'\, p(z')^l$$
$$= \int_0^1 \mathrm{d}u\Big(\int_{-M+u}^{-M} \mathrm{d}z' p(z')^l + \int_{-M}^{M} \mathrm{d}z' p(z')^l + \int_{M}^{M+u} \mathrm{d}z' p(z')^l\Big),$$

where (a) is obtained by first changing the order of integration (which is admissible since the integral converges) and then making the change of variable $z' = z + u$. And so, by combining this identity with (13) we obtain

$$\int_{-M}^{M} \mathrm{d}z \left\{ p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u)\Big)^l \right\}$$
$$+ \int_0^1 \mathrm{d}u \int_{-M+u}^{-M} \mathrm{d}z'\, p(z')^l + \int_0^1 \mathrm{d}u \int_{M}^{M+u} \mathrm{d}z'\, p(z')^l \geq 0$$

Now we take the limit $M \to +\infty$ for each term of this inequality. By an application of Lebesgue's dominated convergence theorem, the last two terms tend to zero and $\frac{1}{2}p^l_{\text{MAP}}$, respectively. Therefore the limit of the first term is bounded from below by $-\frac{1}{2}p^l_{\text{MAP}}$, which concludes the proof. ■

B. Truncation of Profiles

Proof of Lemma 4.2: It is easy to prove that a truncation of $p(z)$ at p_{MAP} yields a smaller value for the single system potential $\mathcal{W}_s(p(z))$ (see e.g. the Figure I for an intuition). Therefore we have $\mathcal{W}_{\text{single}}[p(\cdot)] \geq \mathcal{W}_{\text{single}}[\bar{p}(\cdot)]$.

We now treat the functional corresponding to the interaction term. We define the function g as $g(z) = p(z) - \bar{p}(z)$ and notice that:

$$\begin{cases} p(z) \leq p_{\text{MAP}} \Rightarrow g(z) = 0 \text{ and } \bar{p}(z) = p(z), \\ p(z) > p_{\text{MAP}} \Rightarrow g(z) > 0 \text{ and } \bar{p}(z) = p_{\text{MAP}}. \end{cases} \tag{14}$$

We need to show that $\mathcal{W}_{\text{int}}[\bar{p}(\cdot)] \leq \mathcal{W}_{\text{int}}[p(\cdot)]$, or equivalently that:

$$\int_{\mathbb{R}} \mathrm{d}z \left\{ \bar{p}(z)^l - \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^l \right\}$$
$$\leq \int_{\mathbb{R}} \mathrm{d}z \left\{ (\bar{p}(z) + g(z))^l - \Big(\int_0^1 \mathrm{d}u\, (\bar{p}(z+u) + g(z+u))\Big)^l \right\}$$

Using the binomial expansion this is equivalent to

$$\sum_{i=0}^{l-1} \binom{l}{i} \int_{\mathbb{R}} \mathrm{d}z \left\{ \bar{p}(z)^i g(z)^{l-i} \right.$$
$$\left. - \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i} \right\} \geq 0$$

In the following steps, we show that the integral inside the summation above is positive for any fixed value of i; the inequality follows directly. We see that:

$$\Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i}$$
$$\leq p^i_{\text{MAP}} \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i} \leq p^i_{\text{MAP}} \int_0^1 \mathrm{d}u\, g(z+u)^{l-i},$$

where the first inequality is due to the property $\bar{p}(z) \leq p_{\text{MAP}}$ and the second is using the convexity of the function $f(g) = g^l;\ g \geq 0$. We integrate over z, then make the change of variable $z' = z + u$ on the right-hand side to obtain:

$$\int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i}$$
$$\leq \int_{\mathbb{R}} \mathrm{d}z\, p^i_{\text{MAP}} \int_0^1 \mathrm{d}u\, g(z+u)^{l-i} = \int_{\mathbb{R}} \mathrm{d}z'\, p^i_{\text{MAP}} g(z')^{l-i} \tag{15}$$

Using the properties of g in (14), we remark that $\int_{\mathbb{R}} \mathrm{d}z\, p^i_{\text{MAP}} g(z)^{l-i} = \int_{\mathbb{R}} \mathrm{d}z\, \bar{p}(z)^i\, g(z)^{l-i}$ and so the difference of quantities in the inequality (15) is integrable, and thus we obtain:

$$\int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i} \leq \int_{\mathbb{R}} \mathrm{d}z\, \bar{p}(z)^i g(z)^{l-i}$$

for any i. This yields the desired result $\mathcal{W}_{\text{int}}[p(\cdot)] \geq \mathcal{W}_{\text{int}}[\bar{p}(\cdot)]$. ■

C. Rearrangement of Profiles

Before proceeding with the proof of Lemma 4.3, we state a general rearrangement inequality of Brascamp, Lieb and Luttinger [12].

Theorem A.1: Let $f_j,\ 1 < j < k$ be nonnegative measurable functions on $\mathbb{R}$, and let $a_{jm},\ 1 < j < k,\ 1 < m < n$, be real numbers. Then, if f^* is the symmetric decreasing rearrangement of f, we have:

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \prod_{j=1}^{k} f_j\left(\sum_{m=1}^{n} a_{jm} x_m\right) \leq \int_{\mathbb{R}^n} \mathrm{d}^n x \prod_{j=1}^{k} f^*_j\left(\sum_{m=1}^{n} a_{jm} x_m\right) \tag{16}$$

Remark Theorem A.1 is nontrivial only if $k > n$. Otherwise, both integrals diverge and the inequality trivially holds. We will see in this section that $k > n$ in our case.

Proof of Lemma 4.3: It is sufficient to prove that the increasing rearrangement of a profile decreases $\mathcal{W}_{\rm int}[p(\cdot)]$, since $\mathcal{W}_{\rm single}[p(\cdot)]$ is invariant under any rearrangement.

Theorem A.1 applies to symmetric decreasing rearrangements. Therefore it is convenient to first "symmetrize" the profile and the functional. Consider a profile $p(\cdot) \in \mathcal{S}$ such that $p(z) \in [0, p_{\rm MAP}]$ (due to Lemma 4.2) and denote by $\hat{p}(\cdot)$ the function such that $\hat{p}(z) = p(z), \ z < R$ and $\hat{p}(z) = \hat{p}(2R - z), \ z > R$. The value R is chosen (large enough) so that $p(R)$ is arbitrarily close to $p_{\rm MAP}$. Note that $\hat{p}(\cdot)$ is integrable over $\mathbb{R}$.

We recall the expression of $\mathcal{W}_{int}[p(\cdot)]$ in (5) and rewrite it as:

$$\mathcal{W}_{\rm int}[p(\cdot)] = \frac{\epsilon}{l} \lim_{R\to+\infty} \Big\{ \int_{-\infty}^{R} {\rm d}z p(z)^l - \int_{-\infty}^{R} {\rm d}z \Big(\int_0^1 {\rm d}u \ p(z+u)\Big)^l \Big\}.$$

We now express both integrals in the bracket in terms of the symmetrized profile. For the first one, this is immediate

$$\int_{-\infty}^{R} {\rm d}z p(z)^l = \frac{1}{2}\int_{\mathbb{R}} {\rm d}z \hat{p}(z)^l. \tag{17}$$

For the second one, some care has to be taken with the averaging over u when z is near R. One has

$$\int_{-\infty}^{R} {\rm d}z \Big(\int_0^1 {\rm d}u \ p(z+u)\Big)^l = \frac{1}{2}\int_{\mathbb{R}} {\rm d}z \Big(\int_0^1 {\rm d}u \ \hat{p}(z+u)\Big)^l + o(\frac{1}{R^l}). \tag{18}$$

Replacing these two formulas in (17) we have the representation

$$\mathcal{W}_{\rm int}[p(\cdot)] = \frac{\epsilon}{2l}\int_{\mathbb{R}} {\rm d}z \hat{p}(z)^l - \frac{\epsilon}{2l}\int_{\mathbb{R}} {\rm d}z \Big(\int_0^1 {\rm d}u \ \hat{p}(z+u)\Big)^l. \tag{19}$$

Now consider $\hat{p}^*(\cdot)$, the symmetric decreasing rearrangement of $\hat{p}(\cdot)$. The first term in (19) is invariant under rearrangement. It remains to prove that the second term in (19) increases upon rearrangement. We express it as follows (dropping $\epsilon/2l$):

$$\begin{aligned}
&\int_{\mathbb{R}} {\rm d}z \ \Big(\int_0^1 {\rm d}u \ \hat{p}(z+u)\Big)^l \\
&= \int_{\mathbb{R}} {\rm d}z \int_{\mathbb{R}^l} \prod_{i=1}^{l} {\rm d}u_i \ \hat{p}(z+u_i) \mathbb{1}_{[0,1]}(u_i) \\
&\overset{(b)}{=} \int_{\mathbb{R}} {\rm d}z' \int_{\mathbb{R}^l} \prod_{i=1}^{l} {\rm d}u'_i \ \hat{p}(z'+R+u'_i+\frac{1}{2}) \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u'_i) \\
&\overset{(c)}{\geq} \int_{\mathbb{R}} {\rm d}z' \int_{\mathbb{R}^l} \prod_{i=1}^{l} {\rm d}u'_i \ \hat{p}^*(z'+R+u'_i+\frac{1}{2}) \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u'_i) \\
&\overset{(d)}{=} \int_{\mathbb{R}} {\rm d}z \int_{\mathbb{R}^l} \prod_{i=1}^{l} {\rm d}u_i \ \hat{p}^*(z+u_i) \mathbb{1}_{[0,1]}(u_i) \\
&= \int_{\mathbb{R}} {\rm d}z \ \Big(\int_0^1 {\rm d}u \ \hat{p}^*(z+u)\Big)^l,
\end{aligned}$$

where the equality in (b) is due to the changes of variables $z' = z - R$ and $u'_i = u_i - \frac{1}{2}; \ i = 1 \ldots l$, the inequality in (c) is due Theorem A.1, and the equality in (d) is obtained by first remarking that the indicator function $\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u'_i)$ is unchanged upon rearrangement and then by making the reverse changes of variables $z = z' + R$ and $u_i = u'_i + \frac{1}{2}; \ i = 1 \ldots l$. So far we have obtained

$$\mathcal{W}_{\rm int}[p(\cdot)] \geq \frac{\epsilon}{2l}\int_{\mathbb{R}} {\rm d}z \hat{p}^*(z)^l - \frac{\epsilon}{2l}\int_{\mathbb{R}} {\rm d}z \Big(\int_0^1 {\rm d}u \ \hat{p}^*(z+u)\Big)^l.$$

To obtain $\mathcal{W}_{\rm int}[p(\cdot)] \geq \mathcal{W}_{\rm int}[p^*(\cdot)]$ it remains to reverse the steps (17)-(19). ∎

D. Tightness of the minimizing sequence

Proof of Lemma 4.4: Consider a minimizing sequence of cdfs $p_n(\cdot)$, i.e. satisfying (8). Fix any $\delta > 0$ and suppose that

$$p_n(M) - p_n(-M) < (1-\delta)p_{\rm MAP}. \tag{20}$$

We will show that (20) implies that necessarily $M \leq c/\delta^2$ for a fixed constant $c > 0$. Taking the contrapositive we find that: choosing $M_\delta = c'/\delta^2$ with $c' > c$ implies that any minimizing sequence satisfies $p_n(M_\delta) - p_n(-M_\delta) > (1-\delta)p_{\rm MAP}$.

From Lemma 4.1 we have

$$\mathcal{W}[p_n(\cdot)] \geq \mathcal{W}_{\rm single}[p_n(\cdot)] - \frac{\epsilon p_{\rm MAP}^l}{2l} \geq \int_{-M}^{M} dz \, W_s(p_n(z)) - \frac{\epsilon p_{\rm MAP}^l}{2l}. \tag{21}$$

Now, assuming (20) there must be a mass at least $\delta p_{\rm MAP}$ outside of the interval $[-M, M]$. Thus we have $p_n(-M) \geq \delta p_{\rm MAP}/2$ or $p_{\rm MAP} - p_n(M) \geq \delta p_{\rm MAP}/2$. So in $[-M, M]$ the profile $p_n(z)$ is $\delta p_{\rm MAP}/2$ away from the minima 0 and $p_{\rm MAP}$. Moreover, one can check that $W_s(p)$ has a parabolic shape near the minima at 0 and $p_{\rm MAP}$ so that away from these minima $W_s(p) \geq C\delta^2 p_{\rm MAP}^2/4$

for a constant $C > 0$ depending only on l. These remarks imply

$$\int_{-M}^{M} dz\, W_s(p_n(z)) \geq \frac{1}{2} M C \delta^2 p_{\text{MAP}}^2. \tag{22}$$

Since $p_n(\cdot)$ is a minimizing sequence, for n large enough its cost must be smaller than the cost of a fixed reference profile, say $\rho(z) = 0, z \leq 0$, $\rho(z) = p_{\text{MAP}}, z > 0$. More formally,

$$\mathcal{W}[p_n(\cdot)] < \mathcal{W}[\rho(\cdot)] = -\frac{\epsilon}{l(l+1)} p_{\text{MAP}}^l. \tag{23}$$

Finally, combining (21), (22) and (23) we find that

$$M \leq \frac{\epsilon(l-1)}{l(l+1)} \frac{p_{\text{MAP}}^{l-2}}{C\delta^2}. \tag{24}$$

■

E. Explicit expressions of Kernel Function

In this section, we compute the kernel function V of section V-B and illustrate some of its properties. In particular we show that it is strictly convex in a non-zero measure set.

Recall that the function is totally symmetric under permutations, and it is therefore enough to compute it in a fixed sector $S_x = \{\mathbf{x} = (x_1, \cdots, x_l) : x_i \geq x_j \text{ if } i < j\}$. We express V in terms of the distances d_{1i}, which are ordered such that $d_{1i} < d_{1j}$ if $i < j$.

Let us first discuss the explicit examples $l = 2$ and $l = 3$. For $l = 2$ an explicit computation yields,

$$V_{(l=2)}(d_{12}) = \begin{cases} -\frac{1}{2} & \text{if} \quad d_{12} \geq 1, \\ -\frac{1}{2} + \frac{1}{6}(1 - d_{12})^3 & \text{if} \quad d_{12} < 1. \end{cases}$$

By taking the second derivative it is easy to see that $V_{(l=2)}$ is convex everywhere, and strictly convex for $d_{12} < 1$.

For $l = 3$, we have $d_{12} < d_{13}$ and the computation yields,

$$V_{(l=3)}(d_{12}, d_{13}) = \begin{cases} V_{(l=2)}(d_{12}) & \text{if } d_{13} \geq 1, \\ \\ -\frac{1}{2} + \frac{1}{6}(1 - d_{12})^3 & \\ +\frac{1}{12}(1 - d_{13})^4 & \text{if } d_{13} < 1. \\ +\frac{1}{6} d_{12}(1 - d_{13})^3 & \end{cases}$$

For $d_{13} < 1$ the Hessian is

$$\begin{pmatrix} 1 - d_{12} & -\frac{1}{2}(d_{13} - 1)^2 \\ -\frac{1}{2}(d_{13} - 1)^2 & -(d_{13} - 1)(1 + d_{12} - d_{13}) \end{pmatrix}$$

and the corresponding eigenvalues are:

$$\lambda_{1,2} = \frac{1}{2}\{2 - 2d_{13} - d_{12}d_{13} + d_{13}^2 \pm \sqrt{\Delta}\},$$

where

$$\Delta = 1 + 4d_{12}^2 - 4d_{13} - 8d_{12}d_{13} - 4d_{12}^2 d_{13} + 10 d_{13}^2 + 8 d_{12} d_{13}^2 + d_{12}^2 d_{13}^2 - 8 d_{13}^3 - 2 d_{12} d_{13}^3 + 2 d_{13}^4.$$

A plot of the eigenvalues shows that they are non-negative in the region $0 \leq d_{12} \leq d_{13} \leq 1$. In fact, one eigenvalue is strictly positive everywhere in this region, and the other is strictly positive everywhere in this region except at the boundary $d_{13} = 1$, where it becomes equal to zero. This is consistent with the fact that $V_{(l=3)}(d_{12}, d_{13}) = V_{(l=2)}(d_{12})$ when $d_{13} \geq 1$. For $d_{13} \geq 1$, the Hessian always has a vanishing eigenvalue, and a strictly positive one when $d_{12} < 1$. For $d_{12} \geq 1$, the kernel $V_{(l=3)}(d_{12}, d_{13})$ is constant and both eigenvalues vanish. To summarize the kernel is always convex, and strictly convex for $0 \leq d_{13} < 1$.

These results can be generalized for all l. We find the general expression of the kernel

$$V_l(d_{12}, \cdots, d_{1l}) = \sum_{k=2}^{l} \sum_{m=k}^{l} \frac{(1 - d_{1m})^{m-k+3}}{(m-k+3)(m-k+2)} \times \Big(\sum_{\mathcal{S} \subseteq \{2,\ldots,m-1\}} \prod_{n \in \mathcal{S}} d_{1n} \Big).$$

The corresponding Hessian (H_{ij}) is a symmetric matrix of dimension $(l-1) \times (l-1)$ with matrix elements that are polynomials in d_{1i}, $i = 1, \ldots, l$. In particular, at $d_{1i} = 0$ for all i we have $H_{ii} = 1$ and $H_{ij} = -\frac{1}{j+1} + \sum_{m=j+1}^{l} \frac{1}{(m-1)(m-2)} = -\frac{1}{l-1}$; $j > i$. Defining $\mathbf{v}$ as the $(l-1)$-dimensional vector of 1's and denoting by $\mathbb{1}$ the $(l-1)$-dimensional identity matrix, we remark that H at the origin can be expressed as

$$H = (1 + \frac{1}{l-1})\mathbb{1} - \frac{1}{l-1}\mathbf{v}\mathbf{v}^{\mathrm{T}}.$$

The eigenvalues of this matrix are $1 + \frac{1}{l-1}$ and 1 (with $1 + \frac{1}{l-1}$ having degeneracy $l-2$). Since these eigenvalues are strictly positive, the Hessian is strictly positive definite at the origin, and thus (by continuity) also in a small neighborhood of the origin. Thus V is a strictly convex function of d_{1i}, $i = 2, \ldots, l$ in a small neighborhood of the origin.

Acknowledgment. We thank Marc Vuffray for discussions on rearrangement inequalities and for bringing reference [12] to our attention.

References

[1] A.J. Felstrom and K.S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," IEEE Trans. Inform. Theory, vol. 45, no. 5, pp. 2181-2190, Sept. 1999.

[2] S. Kudekar, T. Richardson, R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC", in *Information Theory Proceedings (ISIT)*, 2010, pp. 684-688.

[3] A. Yedla, Y. Jian, P. Nguyen, and H. Pfister, "A simple proof of threshold saturation for coupled scalar recursions," 2012, e-print: http://arxiv.org/abs/1204.5703.

[4] S. Kudekar, T. Richardson, R. Urbanke, "Wave-Like Solutions of General One-Dimensional Spatially Coupled Systems," 2012, e-print: http://arxiv.org/abs/1208.5273.

[5] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," 2012, e-print: http://arxiv.org/abs/1201.2999.

[6] S. Kumar, A.J. Young, N. Macris, H.D. Pfister, "A Proof of Threshold Saturation for Spatially-Coupled LDPC Codes on BMS Channels," in *Allerton*, Monticello, Il, 2012.

[7] R. McCann, "A convexity principle for interacting gases," Adv. Math. 128(1), 1997, pp. 153-179.

[8] C. Villani, "Displacement Interpolation and Displacement Convexity," in *Topics in Optimal Transportation*, Rhode Island: American Mathematical Society, 2003, ch.5, pp. 143-165.

[9] G. Alberti and G. Bellettini, "A nonlocal anisotropic model for phase transitions Part I: The optimal profile problem," Math. Annalen. 310, 1998, pp. 527-560.

[10] E.A. Carlen, M.C. Carvalho, R. Esposito, J.L. Lebowitz, and R. Marra, "Displacement Convexity and Minimal Fronts at Phase Boundaries," Arch. Rational Mech. Anal. 194, 2009, pp. 823-847.

[11] S.H. Hassani, N. Macris, R. Urbanke, "Chains of Mean-Field Models," Journal of Statistical Mechanics: Theory and Experiment 2012 (02), P02011.

[12] H.J. Brascamp, E.H. Lieb, and J.M. Luttinger. "A general rearrangement inequality for multiple integrals," J. Funct. Anal., vo. 17, 1974, pp. 227-237.

[13] B. Dacorogna, *Direct Methods in the Calculus of Variations*, 2nd ed. New York: Springer-Verlag, 1992.

[14] I. Fonseca, G. Leoni (2007), *Modern Methods in the Calculus of Variations: Lp Spaces*, 1st ed. New York: Springer, 2007.